Journal of Water and Soil
https://jsw.um.ac.ir

Research Article
Vol. 38, No. 2, May-June 2024, p. 269-283

Modeling Soil Penetration Resistance Using Regression, Artificial Neural Network and Gene Expression Programming

Sh. Asghari[1*], M. Hasanpour Kashani[2], H. Shahab Arkhazloo[3]

1 and 3- Professor and Associate Professor, Department of Soil Sciences and Engineering, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran, respectively.
(*- Corresponding Author Email: shasghari@uma.ac.ir)
2- Assistant Professor, Water Engineering Department, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran

Received: 12-02-2024 Revised: 05-03-2024 Accepted: 19-03-2024 Available Online: 19-03-2024	**How to cite this article:** Asghari, Sh., Hasanpour Kashani, M., & Shahab Arkhazloo, H. (2024). Modeling soil penetration resistance using regression, artificial neural network and gene expression programming. *Journal of Water and Soil, 38*(2), 269-283. (In Persian with English abstract). https://doi.org/ 10.22067/JSW.2024.86792.1385

Introduction

The penetration resistance (PR) of the soil shows the mechanical resistance of the soil against the penetration of a conical or flat probe; it is important in terms of seed germination, root growth and tillage operations. In general, if the PR value of a soil exceeds 2.5 MPa, the growth and expansion of roots in the soil will be significantly limited. The direct measurement of PR is also a laborious and costly task due to instrumental errors. Therefore, it is useful the use of different models such as multiple linear regression (MLR), artificial neural network (ANN) and gene expression programming (GEP) to estimate PR through easily accessible and low-cost soil characteristics. The objectives of this research were: (1) to obtain MLR, ANN and GEP models for estimating PR from the easily accessible soil variables in forest, range and cultivated lands of Fandoghloo region of Ardabil province, (2) to compare the accuracy of the aforementioned models in estimating soil PR using the coefficient of determination (R^2), root mean square error (RMSE), mean error (ME) and Nash-Sutcliffe coefficient (NS) criteria.

Materials and Methods

Disturbed and undisturbed samples (n = 80) were nearly systematically taken from 0-10 cm soil depth with nearly 50 m distance in forest (n = 20), range (n = 23) and cultivated (n = 37) lands of Fandoghloo region of Ardabil province, Iran (lat. 38° 24' 10" to 38° 24' 25" N, long. 48° 32' 45" to 48° 33' 5" E) in summer 2023. The contents of sand, silt, clay, $CaCO_3$, pH, EC, bulk (BD) and particle density (PD), organic carbon (OC), gravimetric field water content (FWC), mean weight diameter (MWD) and geometric mean diameter (GMD) were measured in the laboratory. Relative bulk density (BDrel) was calculated using BD and clay data. Mean geometric diameter (d_g) and geometric standard deviation (σ_g) of soil particles were computed by sand, silt and clay percentages. The penetration resistance (PR) of the soil was measured in situ using cone penetrometer (analog model) at 5 replicates. Data randomly were divided in two series as 60 data for training and 20 data for testing of models. The SPSS 22 software with stepwise method, MATLAB and Gene Xpro Tools 4.0 software were used to derive multiple linear regression (MLR), artificial neural network (ANN) and gene expression programming (GEP) models, respectively. A feed forward three-layer (2, 5 and 6 neurons in hidden layer) perceptron network and the tangent sigmoid transfer function were used for the ANN modeling. A set of optimal parameters were chosen before developing a best GEP model. The number of chromosomes and genes, head size and linking function were selected by the trial and error

https://doi.org/ 10.22067/JSW.2024.86792.1385

method, as they are 30, 3, 8, and +, respectively. The rates of genetic operators were chosen according to literature studies. The accuracy of MLR, ANN and GEP models in estimating PR were evaluated by coefficient of determination (R^2), root mean square error (RMSE), mean error (ME) and Nash-Sutcliffe coefficient (NS) statistics.

Results and Discussion

The studied soils had clay loam (n = 11), sandy clay loam (n = 6), sandy loam (n = 12), loam (n = 13), silty clay loam (n = 14), silty clay (n = 1) and silt loam (n = 23) textural classes. The values of sand (13.14 to 64.79 %), silt (21.11 to 74.96 %), clay (2.95 to 42.18 %), OC (1.01 to 7.17 %), FWC (11.58 to 50.47 mass percent), BD (0.84 to 1.43 g cm^{-3}) and PR (1.03 to 5.83 MPa) showed good variations in the soils of the studied region. There were found significant correlations between PR with FWC (r = - 0.45^{**}), silt (r = - 0.36^{**}) and σ_g (r = 0.36^{**}). Due to the multicollinearity of silt with σ_g (r = -0.84^{**}), the σ_g was not used as an input variable to estimate PR. Generally, 3 MLR, ANN and GEP models were constructed to estimate PR from measured readily available soil variables. The results of MLR, ANN and GEP models showed that the most suitable variables to estimate PR were FWC, silt and BDrel. The values of R^2, RMSE, ME and NS criteria were obtained equal 0.44, 1.19 MPa, 0.19 MPa and 0.36, and 0.92, 0.41 MPa, -0.05 MPa and 0.92, 0.79, 0.91 MPa, 0.13 MPa, 0.63 for the best MLR, ANN and GEP models, respectively. The former researchers also reported that there is a negative and significant correlation between PR with FWC.

Conclusion

The results indicated that field water content (FWC), silt and relative bulk density (BDrel) were the most important and readily available soil variables to estimate penetration resistance (PR) in the studied area. According to the lowest values of RMSE and the highest values of NS, the accuracy of ANN models to predict soil PR was higher than MLR and GEP models in this research.

Keywords: Estimation, Fandoghloo region, Gravitational water content, Readily variables, Soil resistance

نشریه آب و خاک

https://jsw.um.ac.ir

مقاله پژوهشی

جلد ۳۸، شماره ۲، خرداد-تیر ۱۴۰۳، ص. ۲۸۳-۲۶۹

مدل‌سازی مقاومت فروروی خاک با استفاده از رگرسیون، شبکه عصبی مصنوعی و برنامه‌ریزی بیان ژن

شکراله اصغری[۱*]- مهسا حسنپور کاشانی[۲]- حسین شهاب آرخازلو[۳]

تاریخ دریافت: ۱۴۰۲/۱۱/۲۳

تاریخ پذیرش: ۱۴۰۲/۱۲/۲۹

چکیده

اطلاع از مقاومت فروروی (PR) خاک از نظر جوانه‌زنی بذر، رشد ریشه و عملیات خاک‌ورزی اهمیت فراوان دارد. اندازه‌گیری مستقیم PR خاک به‌دلیل تغییرپذیری مکانی و زمانی شدید آن، کاری پرزحمت و گران می‌باشد. هدف از این پژوهش، ارایه مدل‌های رگرسیون خطی (MLR)، شبکه عصبی مصنوعی (ANN) و برنامه‌ریزی بیان ژن (GEP) برای برآورد PR از روی ویژگی‌های زودیافت خاک بود. در مجموع ۸۰ نمونه خاک سطحی (۰ تا cm ۱۰) دست‌خورده و دست‌نخورده (با استفاده از استوانه‌های استیل به قطر و ارتفاع ۵ سانتی‌متر) از اراضی جنگلی، مرتعی و زراعی منطقه فندقلوی اردبیل در تابستان ۱۴۰۲ برداشته شد سپس برخی خصوصیات فیزیکی و شیمیایی زودیافت خاک در آنها اندازه‌گیری شد. مقاومت فروروی خاک به‌طور درجا در محل با استفاده از یک فروسنج مخروطی تعیین گردید و همزمان رطوبت خاک مزرعه در استوانه‌ها اندازه‌گیری شد. داده‌ها به‌طور تصادفی به دو گروه آموزشی (۶۰ نمونه) و آزمونی (۲۰ نمونه) تقسیم گردید. مدل‌های MLR، ANN و GEP به‌ترتیب با استفاده از نرم‌افزارهای SPSS، MATLAB و Gene Xpro Tools ایجاد شدند. نتایج مدل‌سازی نشان داد که رطوبت خاک مزرعه، سیلت و جرم مخصوص ظاهری نسبی، مهمترین متغیرهای ورودی در برآورد PR خاک بودند. مقادیر آماره‌های ضریب تبیین (R^2)، مجذور میانگین مربعات خطا (RMSE)، میانگین خطا (ME) و ضریب نش–ساتکلیف (NS) براساس داده‌های آزمونی برابر ۰/۴۴، MPa ۱/۱۹، MPa ۰/۱۹ و ۰/۳۶، ۰/۹۲، MPa ۰/۴۱، MPa ۰/۰۵– و ۰/۹۲، ۰/۷۹، MPa ۰/۹۱، MPa ۰/۱۳ و ۰/۶۳ به‌ترتیب برای بهترین مدل MLR، ANN و GEP تعیین گردید. براساس نتایج آماره‌های ارزیابی مدل‌ها، می‌توان گفت که در منطقه مورد مطالعه، مدل ANN از بیشترین دقت و مدل MLR از کمترین دقت در برآورد PR خاک برخوردار بود.

واژه‌های کلیدی: تخمین، رطوبت جرمی، مقاومت خاک، منطقه فندقلو، ویژگی زودیافت

مقدمه

مقاومت فروروی[۴] (PR) خاک که مقاومت مکانیکی خاک را در برابر نفوذ یک فروسنج مخروطی یا مسطح نشان می‌دهد، از نظر جوانه زنی بذر، رشد ریشه و عملیات خاک‌ورزی حائز اهمیت است. به‌طور کلی، اگر مقدار PR یک خاک از ۲/۵ مگاپاسکال فراتر رود رشد و گسترش ریشه در خاک به‌طور معنی‌دار محدود خواهد شد (Whalley *et al.*, 2007).

مقاومت فروروی (PR) خاک با تغییر سایر ویژگی‌های خاک مانند کربن آلی، جرم مخصوص ظاهری و رطوبت خاک مزرعه، با زمان و مکان تغییر پیدا می‌کند. به‌طور مثال، اصغری و همکاران (Asghari *et al.*, 2017) گزارش کردند ضریب تغییرات (CV) مقاومت فروروی

۱ و ۳- به‌ترتیب استاد و دانشیار، گروه علوم و مهندسی خاک، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران

(*- نویسنده مسئول: Email: shasghari@uma.ac.ir)

۲- استادیار، گروه مهندسی آب، دانشکده کشاورزی و منابع طبیعی، دانشگاه محقق اردبیلی، اردبیل، ایران

https://doi.org/ 10.22067/JSW.2024.86792.1385

4- Penetration resistance

در ۱۰۵ نمونه خاک برداشته‌شده (با فواصل ۳۰۰۰ متر) از اراضی زراعی دشت اردبیل (۲۴ km × ۴۲) حدود ۵۰ درصد تعیین گردید و دامنه تغییرات کربن آلی از ۰/۳ تا ۲/۴ درصد، جرم مخصوص ظاهری از ۱/۰۲ تا ۱/۶۳ گرم بر سانتی‌متر مکعب و PR از ۱/۱ تا ۶/۶ مگاپاسکال به دست آمد. همچنین اصغری و شهابی (Asghari & Shahabi, 2019) مقادیر CV را برای مقاومت فروروی خاک در اراضی زراعی و بایر بخش شندآباد منطقه شبستر (۲۰۰۰ × ۴۰۰ m) حدود ۴۶ درصد گزارش نمودند. از سوی دیگر، اندازه‌گیری مستقیم PR نیز به‌دلیل خطاهای ابزاری، کاری طاقت‌فرسا و پرهزینه است؛ بنابراین استفاده از مدل‌های مختلف مانند رگرسیون خطی چندگانه[1](MLR)، شبکه عصبی مصنوعی (ANN)[2] و برنامه‌ریزی بیان ژن[3](GEP) برای برآورد PR از طریق ویژگی‌های سهل‌الوصول و کم‌هزینه خاک می‌تواند در این زمینه راهگشا باشد.

نتایج پژوهش بیات و همکاران (Bayat *et al.*, 2008) نشان داد که در برآورد PR از روی تخلخل کل، درجه اشباع نسبی و رطوبت خاک مزرعه، مدل ANN در مقایسه با مدل‌های رگرسیون خطی و غیرخطی از دقت بالاتری برخوردار بود. واز و همکاران (Vaz *et al.*, 2011) گزارش کردند که رطوبت خاک، مهمترین ویژگی تأثیرگذار بر مقاومت فروروی خاک می‌باشد. سانتوس و همکاران (Santos *et al.*, 2011) از جرم مخصوص ظاهری و رطوبت خاک مزرعه برای برآورد PR خاک در مدل‌های رگرسیونی و ANN استفاده کردند. آنها بیان کردند که PR با جرم مخصوص ظاهری رابطه مستقیم و با رطوبت خاک مزرعه رابطه معکوس دارد. شیخزاده و همکاران (Sheykhzadeh *et al.*, 2016) گزارش کردند که در برآورد PR خاک، مقادیر ضریب تبیین (R^2)[4] و مجذور میانگین مربعات خطا (RMSE)[5] به‌ترتیب در مدل ANN برابر با ۰/۹۱ و ۰/۳۷ MPa و در مدل MLR برابر با ۰/۵۵ و ۰/۸۹ MPa به‌دست آمد بنابراین مدل ANN نسبت به مدل MLR دارای دقت بالاتری بود؛ همچنین رطوبت خاک مزرعه، جرم مخصوص ظاهری و توزیع اندازه ذرات خاک، مهمترین متغیرهای ورودی در برآورد PR خاک تشخیص داده شد. اخیراً از سایر روش‌های هوشمند مانند الگوریتم ژنتیک و مدل GEP در برآورد ویژگی‌های دیریافت خاک استفاده گردیده است؛ احمدی و همکاران (Ahmadi *et al.*, 2016) گزارش کردند که در برآورد هدایت هیدرولیکی اشباع خاک در دشت تبریز، مدل GEP نسبت به مدل رگرسیون رایج به‌علت داشتن R^2 بالا و RMSE پایین دارای دقت بالاتری بود. آزاد مرد و همکاران (Azadmard *et al.*, 2019) بیان کردند مدل ترکیبی ANN با الگوریتم ژنتیک در مقایسه با مدل MLR از دقت به مراتب بالاتری در برآورد هدایت هیدرولیکی نزدیک اشباع خاک در دشت مغان استان اردبیل برخوردار بود. نتایج پژوهش احمدزاده کلیبر و فولاپناه (Ahmadzadeh Kaleibar & Fuladipanah, 2023) در دشت ارسباران واقع در شمال غرب ایران نیز نشان داد که مدل GEP در مقایسه با مدل‌های رگرسیون خطی و غیرخطی از دقت بالاتری در برآورد رطوبت‌های ظرفیت مزرعه و نقطه پژمردگی دائم برخوردار بود. زانگ و زانگ (Zhang & Zhangh, 2024) گزارش کردند که مدل GEP توانست ضریب نفوذپذیری خاک را از روی برخی ویژگی‌های خاک مانند نسبت پوکی، اندازه مؤثر و میانگین اندازه ذره با دقت بالاتری (R^2 برابر با ۰/۹۴ و RMSE برابر با ۰/۰۰۰۴۸) نسبت به سایر مدل‌های هوشمند و روابط ریاضی تخمین بزند.

اگرچه مدل‌های رگرسیونی و شبکه عصبی مصنوعی (ANN) مختلفی برای برآورد PR خاک در نقاط مختلف دنیا ارائه گردیده است ولی کاربرد عملی این مدل‌ها در خاک‌های سایر مناطق جهان با شرایط اقلیمی متنوع، همواره توسط محققان با شک و تردید همراه بوده است. از سوی دیگر، تحقیقات انجام گرفته در زمینه استفاده از مدل برنامه ریزی بیان ژن (GEP) برای برآورد PR خاک انگشت‌شمار است. بررسی منابع نشان داد که تاکنون تحقیقی در زمینه برآورد PR توسط مدل‌های هوشمند مانند ANN و GEP و مقایسه آن با مدل‌های رگرسیونی در جنگل‌های فندقلوی اردبیل اجرا نگردیده است لذا موضوع پژوهش حاضر از این نظر دارای نوآوری است. با توجه به اینکه منطقه فندقلوی استان اردبیل از مناطق گردشگری مهم و پرتردد کشور به‌ویژه در فصل تابستان محسوب می‌شود و تغییر کاربری اراضی جنگلی و مرتعی به زراعی نیز در آن با شدت زیاد در حال وقوع است لذا ارائه مدل‌های هوشمند برای اطلاع از وضعیت فیزیکی و مکانیکی خاک‌های منطقه مذکور حائز اهمیت فراوان است. بنابراین اهداف پژوهش حاضر عبارت بودند از (۱) ارائه مدل‌های MLR، ANN و GEP برای برآورد PR از روی متغیرهای زودیافت در خاک‌های جنگلی، مرتعی و زراعی منطقه فندقلوی استان اردبیل (۲) مقایسه دقت مدل‌های مذکور در برآورد PR خاک با استفاده از معیارهای R^2، RMSE، میانگین خطا (ME)[6] و ضریب نش–ساتکلیف (NS)[7] بود.

مواد و روش‌ها

خاک‌های مورد مطالعه، در اراضی منطقه گردشگری فندقلو واقع در شرق استان اردبیل (مختصات جغرافیایی ″۴۵ ′۳۲ °۴۸ تا ″۵ ′۳۳ °۴۸ طول شرقی و ″۱۰ ′۲۴ °۳۸ تا ″۲۵ ′۲۴ °۳۸ عرض شمالی) قرار

1- Multiple linear regression
2- Artificial neural network
3- Gene expression programming
4- Determination coefficient
5- Root mean square error
6- Mean error
7- Nash-Sutcliffe coefficient

دارد که فاصله آن از مرکز استان اردبیل در حدود ۳۰ کیلومتر می‌باشد. تعداد ۸۰ نمونه خاک دست‌خورده (با حداقل دست‌خوردگی) و دست نخورده (با استفاده از استوانه‌های استیل به قطر و ارتفاع ۵ سانتی‌متر) از محدوده عمق ۰ تا ۱۰ سانتی‌متر اراضی شیبدار و تپه‌ای با کاربری جنگلی (۲۰ نمونه)، مرتعی (۲۳ نمونه) و زراعی (۳۷ نمونه) به هم چسبیده به فواصل حدود ۵۰ متر و به طور تصادفی- منظم در تابستان ۱۴۰۲ برداشته‌شد. موقعیت هر یک از نقاط نمونه‌برداری‌شده با استفاده از دستگاه موقعیت‌یاب جهانی (GPS)[1] ثبت گردید (شکل ۱).

پس از هواخشک کردن نمونه‌های خاک، بخشی از آن‌ها برای تعیین پایداری خاکدانه‌ها جدا گردیده و بخش دیگر برای تعیین سایر ویژگی‌های فیزیکی و شیمیایی از الک ۲ mm گذرانده شد. توزیع اندازه ذرات شن، سیلت و رس و کلاس بافت خاک به روش هیدرومتری۴ قرائته (Gee & Or, 2002) تعیین و پارامترهای میانگین هندسی قطر (d_g) و انحراف استاندارد هندسی (σ_g) ذرات خاک از روی درصدهای شن، سیلت و رس محاسبه گردید (Campbell, 1985). کربن آلی به روش اکسیداسیون تر (Nelson & Sommers, 1982)، اسیدیته (pH) و هدایت الکتریکی (EC) در عصاره ۱ به ۲/۵ خاک به آب و آهک به روش تیتراسیون با اسید و سود (Page, 1985) تعیین شد. جرم مخصوص ظاهری (BD) در استوانه‌های استیل (Blake & Hartge, 1986a) و جرم مخصوص حقیقی (PD) به روش پیکنومتر (Blake & Hartge, 1986b) اندازه‌گیری شد. جرم مخصوص ظاهری نسبی (BDrel) از روی داده‌های جرم مخصوص ظاهری و درصد رس محاسبه گردید (Kelishadi *et al.*, 2014). میانگین وزنی قطر (MWD)[2] و میانگین هندسی قطر (GMD)[3] خاکدانه‌ها به روش الک خشک در ۵۰ گرم خاک هواخشک گذرانده شده از الک ۴/۷۵ میلی‌متر با استفاده از دستگاه شیکر الک با به‌کارگیری الک‌های با قطر سوراخ به‌ترتیب از بالا به پایین ۲، ۱، ۰/۵، ۰/۲۵ و ۰/۱۰۶ میلی‌متر اندازه‌گیری و محاسبه گردید (Yoder, 1936). مقاومت فروروی (PR) خاک به‌طور مستقیم در مزرعه با استفاده از دستگاه فروسنج مخروطی مدل عقربه‌ای ساخت شرکت اکل کمپ کشور هلند در ۵ تکرار اندازه‌گیری شد (Lowery & Morrison, 2002) و همزمان رطوبت خاک مزرعه در حین اندازه‌گیری PR، با استفاده از استوانه‌های خاک برداشته شده برای اندازه‌گیری BD پس از آون خشک کردن در دمای ۱۰۵ C^{o} به روش وزنی تعیین گردید (Gardner, 1986).

نرمال بودن توزیع داده‌ها با آزمون کولموگروف— اسمیرنوف بررسی گردید سپس محاسبات مربوط به آمار توصیفی و ضریب همبستگی پیرسون بین متغیرها و نیز اشتقاق مدل‌های رگرسیونی خطی چندگانه (MLR) بین متغیرهای زودیافت خاک با متغیر دیریافت (PR) به روش گام به گام در نرم‌افزار SPSS 22 انجام شد.

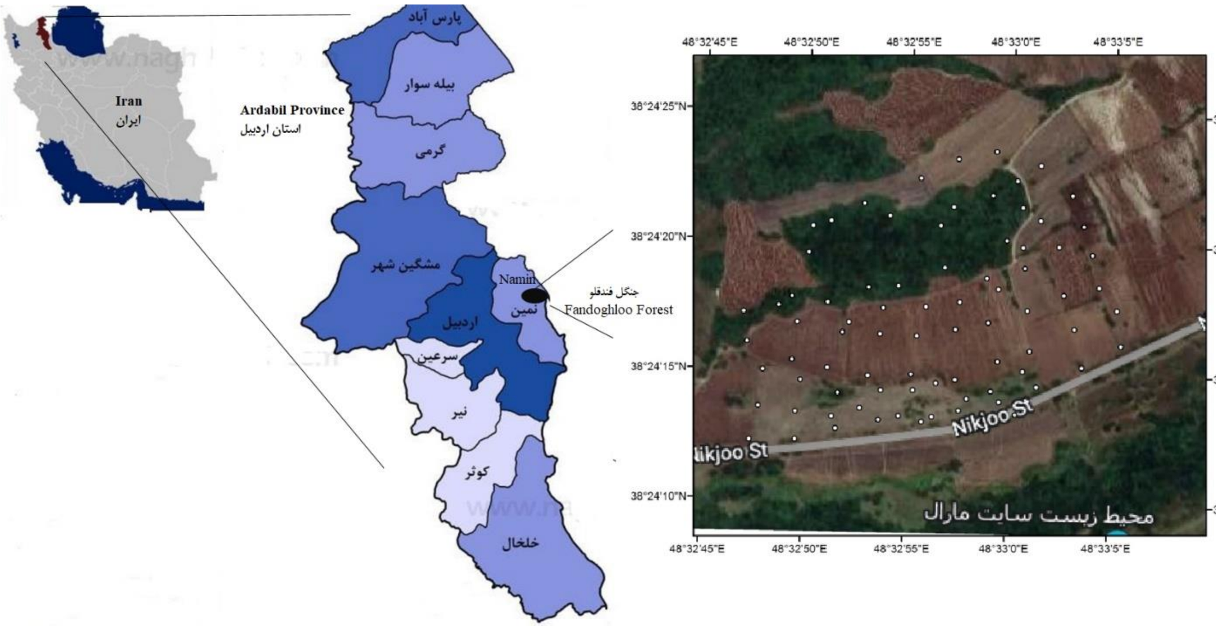
مدل‌های شبکه عصبی مصنوعی (ANN) با استفاده از نرم‌افزار MATLAB ایجاد گردید. در این پژوهش شبکه پرسپترون سه لایه شامل یک لایه ورودی، یک لایه پنهان و یک لایه خروجی با الگوریتم پس انتشار خطا و ۱۰۰۰ تکرار برای رسیدن به حداقل خطا و حداکثر پایداری شبکه به کار گرفته‌شد و تعداد نرون‌های مناسب به روش سعی و خطا در لایه پنهان برای سه مدل شبکه عصبی به‌ترتیب برابر ۶، ۵ و ۲ انتخاب گردید. نرون‌های پنهان داخل شبکه پرسپترون چند لایه، در هر لایه به نرون‌های لایه بعد متصل هستند. به‌منظور تجزیه و تحلیل ویژگی‌ها در داده‌های ورودی، توابع فعال‌سازی تانژانت سیگموئید و خطی به‌ترتیب در لایه‌های پنهان و خروجی هر سه مدل به‌کار برده شد (Ghorbani *et al.*, 2019). به‌منظور آموزش شبکه با دقت بالا از الگوریتم آموزشی مارکوآرت–لونبرگ استفاده گردید.

برنامه‌ریزی بیان ژن (GEP) توسط فریرا (Ferreira, 2001) پیشنهاد شد. این روش یک الگوریتم توسعه‌یافته براساس برنامه‌ریزی ژنتیک و الگوریتم ژنتیک می‌باشد و مبنای تمامی آنها براساس نظریه تکامل داروین استوار است که در آن کروموزوم‌های خطی و ساده با طول ثابت مشابه آنچه در الگوریتم ژنتیک استفاده می‌شود و ساختارهای شاخه‌ای با اندازه‌ها و اشکال متفاوت مشابه با برنامه‌ریزی ژنتیک ترکیب می‌شوند. برای مطالعه بیشتر در زمینه جزئیات مدل GEP می‌توان به مقاله زانگ و زانگ (Zhang & Zhangh, 2024) مراجعه نمود. در این تحقیق برای مدل‌سازی مقاومت فروروی (PR) خاک براساس برنامه ریزی بیان ژن (GEP) از نرم‌افزار Gene Xpro Tools 4.0 استفاده شد. برای این منظور از مجموعه عملگرهای ریاضی شامل چهار عمل اصلی (+، −، × و /)، جذر، اکسپونشیال، لگاریتم طبیعی، توان ۲ و ۳ و ریشه سوم استفاده شد. همچنین پارامترهای لازم برای مدل‌سازی GEP در جدول ۱ آورده شده است.

در همه مدل‌های به‌کار رفته در تحقیق حاضر از مجموع ۸۰ داده برای هر متغیر، بعد از تصادفی کردن کل داده‌ها در نرم‌افزار اکسل، ۶۰ داده برای آموزش مدل و ۲۰ داده برای آزمون مدل در نظر گرفته شد. برای ارزیابی دقت و کارآیی مدل‌های MLR، ANN و GEP در برآورد PR، از محک‌های آماری ضریب تبیین (R^2)، مجذور میانگین مربعات خطا (RMSE)، میانگین خطا (ME) و ضریب نش–ساتکلیف (NS) (Ghorbani et *al.*, 2019) استفاده شد. مقادیر بالای R^2 (نزدیک به ۱) بیانگر دقت بالای برآورد می‌باشد.

3- Geometric mean diameter

1- Global positioning system
2- Mean weight diameter



شکل ۱- موقعیت جغرافیایی و نقشه توزیع نقاط نمونه‌برداری در تصویر ماهواره‌ای منطقه مورد مطالعه

Figure 1- Geographic position and distribution map of sampled points in satellite image of the studied area

جدول ۱- مقادیر پارامترهای مدل برنامه‌ریزی بیان ژن (GEP)

Table 1- The values of GEP model parameters

مقادیر Values	عوامل ژنتیک Genetic factors	مقادیر Values	تنظیمات عمومی General adjustment
0.044	Mutation rate نرخ جهش	30	Number of chromosome تعداد کروموزم‌ها
0.1	Inversion rate نرخ وارون‌سازی	8	Head size اندازه سر
0.3	One point recombination rate نرخ ترکیب تک نقطه‌ای	3	Number of genes تعداد ژن‌ها
0.3	Two point recombination rate نرخ ترکیب دو نقطه‌ای	جمع (+)	Linking function تابع پیوند
0.1	Gene recombination rate نرخ ترکیب ژن	RMSE	Fitness function error type معیارخطای تابع

از لحاظ آماری کمترین مقدار RMSE، صفر می‌باشد که در آن وضعیت، مقادیر اندازه‌گیری و برآورد شده PR باهم برابر هستند؛ هرچه مقدار RMSE کوچکتر باشد برآوردهای انجام شده توسط مدل برآورد کننده دقیق‌تر است. مقادیر مثبت ME بیش‌برآوردی و مقادیر منفی آن کم‌برآوردی مدل را نشان می‌دهد. اگر ضریب NS برابر یک باشد بیانگر آن است که مدل‌سازی با دقت بالایی انجام گرفته و مقادیر اندازه‌گیری و برآورد شده PR در تطابق کامل هستند. مقدار برابر با صفر NS نشان‌دهنده آن است که برآوردهای مدل، دقتی برابر با دقت میانگین داده‌های اندازه‌گیری شده دارند.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_{i(p)} - Y_{i(m)})^2}{n}} \qquad (۱)$$

$$ME = \frac{\sum_{i=1}^{n} Y_{i(p)} - Y_{i(m)}}{n} \qquad (۲)$$

$$NS = 1 - \frac{\sum_{i=1}^{n}(Y_{i(p)} - Y_{i(m)})^2}{\sum_{i=1}^{n}(Y_{i(m)} - \overline{Y_m})^2} \qquad (۳)$$

در معادله‌های ۱، ۲ و ۳، $Y_{i(m)}$، $Y_{i(p)}$ و $\overline{Y_m}$ به‌ترتیب مقادیر اندازه گیری، برآورد شده و میانگین مقادیر اندازه‌گیری شده PR توسط مدل و n تعداد داده‌ها برای PR در هر مدل می‌باشد. از میان مدل‌های ایجاد شده برای برآورد PR، مناسب‌ترین مدل MLR، ANN و GEP با ترکیب متغیرهای ورودی در داده‌های آزمونی براساس محک‌های آماری مذکور تعیین گردید. آزمون مدل‌ها در تشخیص دقت مدل به ازای داده‌هایی که در آموزش مدل استفاده نشده‌اند به کار می‌رود.

نتایج و بحث

آماره‌های توصیفی ویژگی‌های خاک

توصیف آماری ویژگی‌های خاک منطقه فندقلوی اردبیل در جدول ۲ نشان می‌دهد که داده‌های آموزشی و آزمونی به‌کار رفته در توسعه و ارزیابی مدل‌های رگرسیون خطی چندگانه (MLR)، شبکه عصبی مصنوعی (ANN) و برنامه‌ریزی بیان ژن (GEP) دارای توزیع نسبتاً مناسبی هستند و این موضوع از نظر توسعه و ارزیابی مدل‌ها حائز اهمیت است؛ به‌طور مثال، دامنه تغییرات بالا به‌ترتیب در داده‌های آموزشی و آزمونی برای مقاومت فروروی خاک (PR) از ۱/۲ تا ۴/۸ مگاپاسکال و از ۱/۰۳ تا ۵/۸۳ مگاپاسکال و برای رطوبت خاک مزرعه (FWC) از ۱۱/۴۸ تا ۵۰/۴۴ درصد وزنی و ۱۲/۹۰ تا ۵۰/۴۷ درصد وزنی و کربن آلی خاک از ۱/۰۱ تا ۷/۱۷ و ۱/۱۵ تا ۶/۳۱ درصد مؤید این موضوع است. همچنین مقادیر میانگین داده‌ها در سری آموزش و آزمون به هم نزدیک بوده و به‌ترتیب برای PR ۲/۱۵ و ۲/۲۲ مگاپاسکال، برای FWC ۲۴/۴۶ و ۲۵/۷۳ درصد وزنی و برای کربن آلی، ۲/۸۵ و ۲/۹۰ درصد می‌باشد. لازم به توضیح است که مقادیر بالای کربن آلی به کاربری جنگلی و مقادیر پایین کربن آلی به کاربری زراعی مربوط می‌باشد. همچنین مقادیر بالای به‌دست آمده برای رطوبت خاک مزرعه (در لحظه اندازه‌گیری PR) به‌علت وقوع بارندگی در اواخر تیر ۱۴۰۲ در منطقه مورد مطالعه می‌باشد. مقادیر متوسط میانگین وزنی (MWD) و هندسی (GMD) قطر خاکدانه‌ها به‌ترتیب در داده‌های سری آموزش برابر با ۱/۵۵ و mm ۱/۰۳ و در داده‌های سری آزمون برابر با ۱/۶۱ و mm ۱/۴ به‌دست آمد (جدول ۲)؛ مقادیر بالای شاخص‌های پایداری خاکدانه‌ها (MWD و GMD) در تحقیق حاضر در مقایسه با سایر خاک‌های استان اردبیل را می‌توان به کربن آلی زیاد خاک در منطقه مورد مطالعه نسبت داد. شیخ‌زاده و همکاران (Sheykhzadeh *et al.*, 2016) میانگین کربن آلی و PR خاک را به‌ترتیب ۱/۱۱ درصد و ۲/۷۸ مگاپاسگال در ۱۰۵ نمونه خاک زراعی برداشته‌شده از دشت اردبیل گزارش کردند. امیرعابدی و همکاران (Amirabedi *et al.*, 2016) نیز میانگین کربن آلی و MWD را به ترتیب ۰/۷۷ درصد و mm ۰/۵۰ در ۱۰۰ نمونه خاک زراعی برداشته شده از دشت اردبیل گزارش کردند.

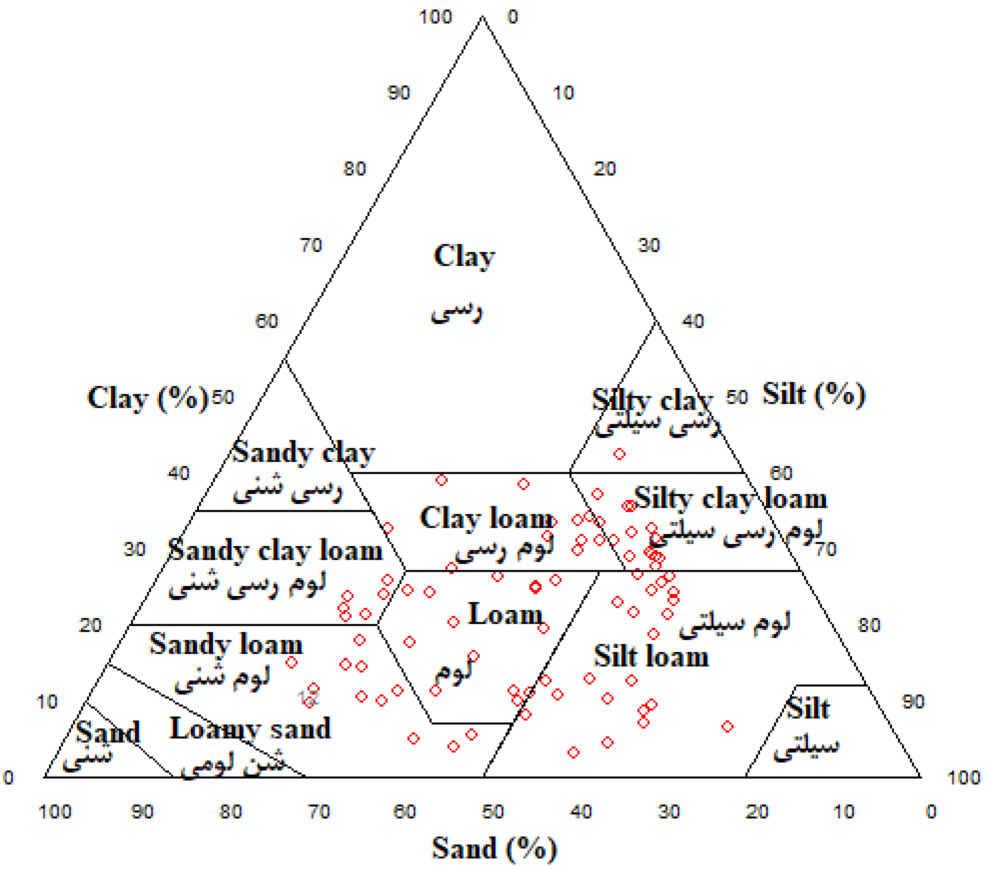
با توجه به شکل ۲ مشاهده می‌شود که کلاس بافت خاک‌های مورد مطالعه براساس مثلث بافت آمریکایی لوم، لوم رسی سیلتی، لوم شنی، لوم رسی، لوم سیلتی و رسی سیلتی بود که بیانگر توزیع گسترده بافتی در منطقه مورد مطالعه است.

جدول ۲- آماره‌های توصیفی متغیرهای خاک در منطقه مورد مطالعه

Table 2- Descriptive statistics of soil variables in the studied area

متغیر Variable	آموزش Training				آزمون Testing			
	Min.	Max.	Mean	Sd.	Min.	Max.	Mean	Sd.
FWC (% g g^{-1})	11.48	50.44	24.46	10.24	12.90	50.47	25.73	12.15
BDrel	0.44	0.75	0.63	0.07	0.43	0.79	0.63	0.10
PD (g cm^{-3})	2.01	2.67	2.44	0.15	2.02	2.68	2.41	0.17
Clay (%)	2.95	42.18	21.78	9.53	3.94	38.40	20.28	11.27
Silt (%)	21.11	65.19	44.62	12.90	25.53	74.96	47.45	12.35
Sand (%)	13.14	64.10	33.59	15.52	15.31	64.80	32.27	14.89
d_g (mm)	0.01	0.18	0.06	0.05	0.02	0.21	0.06	0.05
σ_g (-)	6.94	20.43	11.75	3.29	6.94	20.43	11.75	3.29
MWDd (mm)	0.98	2.31	1.55	0.28	1.19	1.98	1.61	0.25
GMD (mm)	0.8	1.33	1.03	0.11	0.8	1.2	1.04	0.1
OC (%)	1.01	7.17	2.85	1.75	1.15	6.31	2.90	1.71
$CaCO_3$ (%)	3.12	10.32	7.18	1.76	2.1	15.6	6.74	2.67
$pH_{1:2.5}$	5.45	7.2	6.29	0.38	5.68	7.34	6.33	0.43
$EC_{1:2.5}$ (dS m^{-1})	0.02	1.08	0.15	0.19	0.04	0.33	0.11	0.07
PR (MPa)	1.2	4.8	2.15	0.85	1.03	5.83	2.22	1.52

FWC: رطوبت خاک مزرعه (field water content)، BDrel:جرم مخصوص ظاهری نسبی (relative bulk density)، D_p: جرم مخصوص حقیقی (particle density)، Clay: رس، silt: سیلت، sand: شن، d_g: میانگین هندسی قطر ذرات (geometric mean diameter)، σ_g: انحراف استاندارد هندسی ذرات (geometric standard deviation)، MWDd: میانگین وزنی قطر خاکدانه‌ها به روش الک خشک (mean weight diameter)، GMD: میانگین هندسی قطر خاکدانه‌ها (geometric mean diameter). OC: کربن آلی (organic carbon)، $CaCO_3$: کربنات کلسیم معادل، PR: مقاومت فروروی (penetration resistance). Sd: انحراف استاندارد، Mean: میانگین، Max: حداکثر، Min: حداقل.



شکل ۲- توزیع کلاس بافت در خاک‌های مورد مطالعه

Figure 2- Distribution of textural class in the studied soils

ضرایب همبستگی پیرسون (r) بین متغیرهای خاک در جدول ۳ نشان می‌دهد که همبستگی خطی قوی (r >0.60) بین شن با سیلت و میانگین هندسی قطر ذرات (d_g)، سیلت با انحراف استاندارد هندسی ذرات (σ_g)، رس با d_g، کربن آلی با رطوبت خاک مزرعه و MWDd با GMD وجود دارد بنابراین برای اجتناب از همراستایی چندگانه[1] از این متغیرها به‌عنوان ورودی مدل به‌طور همزمان برای ایجاد مدل‌های MLR در برآورد PR استفاده نشد. عامل تورم واریانس[2] (VIF) از جمله روش‌های تشخیص همراستایی چندگانه بوده که نشان می‌دهد واریانس ضرایب برآوردی تا چه حد نسبت به‌حالتی‌که متغیرهای برآوردی همبستگی خطی ندارند افزایش یافته است (Yazdani *et al.*, 2014). نتایج آنالیز هم‌خطی حداکثر تورم واریانس (VIF) مدل‌های MLR مربوطه (جدول ۴) مقادیر کمتر از ۵ را نشان دادند که مقدار حداکثر آن در این پژوهش ۱/۴ به‌دست آمد. به‌طور کلی، اگر آماره VIF برای یک متغیر کمتر از ۵ باشد آن متغیر در مدل وارد می‌گردد در غیر این صورت باید از مدل حذف شود.

نتایج مدل‌های رگرسیون خطی چندگانه (MLR) برای برآورد مقاومت فروروی (PR) خاک

به‌طور کلی براساس روش گام به گام در رگرسیون خطی، ۳ نوع مدل MLR برای برآورد PR با استفاده از متغیرهای رطوبت خاک مزرعه (FWC)، سیلت و جرم مخصوص ظاهری نسبی (BDrel) از بین متغیرهای زودیافت در منطقه مورد مطالعه تشخیص داده شد (جدول ۴). نتایج ضریب همبستگی پیرسون بین PR با متغیرهای زودیافت خاک در جدول ۳ نشان می‌دهد که بیش‌ترین همبستگی منفی و معنی‌دار بین PR با FWC (r = -۰/۴۵**) یافت شد. همچنین همبستگی منفی و معنی‌دار بین PR با سیلت (r = -۰/۳۶**) و همبستگی مثبت و معنی‌دار بین PR با انحراف استاندارد هندسی ذرات (σ_g) (r = ۰/۳۶**) به‌دست آمد؛ البته به‌دلیل همراستایی سیلت و σ_g، از σ_g به‌عنوان ورودی در تخمین PR استفاده نشد و سایر متغیرهای

1- Multicolinearity

2- Variance inflation factor (VIF)

ورودی (به استثنای BDrel) نیز براساس تشخیص روش گام به گام و احتمالاً به علت نداشتن همبستگی معنی‌دار با PR، در مدل‌های MLR ظاهر نشدند.

مشاهده می‌شود FWC در هر سه مدل MLR به‌عنوان متغیر ورودی و با علامت منفی ظاهر گردیده است (جدول ۴) بنابراین در پژوهش حاضر رطوبت خاک مزرعه (FWC) مهمترین عامل اثرگذار در برآورد مقاومت فروروی خاک بوده است. در مدل‌های رگرسیونی ارائه شده توسط بیات و همکاران (Bayat *et al.*, 2008)، واز و همکاران (Vaz *et al.*, 2011) و شیخزاده و همکاران (Sheykhzadeh *et al.*, 2016) نیز رطوبت خاک مزرعه یکی متغیرهای ورودی مهم در برآورد PR گزارش گردیده است. علامت منفی FWC در هر سه مدل MLR نیز بیانگر آن است که با افزایش رطوبت خاک محل و نرم شدن خاک، مقاومت مکانیکی خاک در برابر نفوذ فروسنج مخروطی کاهش یافته است که با یافته‌های سانتوس و همکاران (Santos *et al.*, 2011) و اصغری و شهابی (Asghari & Shahabi, 2019) در یک راستا می باشد. همچنین مشاهده می‌شود در مدل MLR شماره ۲ و ۳، سیلت با علامت منفی به‌عنوان متغیر ورودی وارد مدل گردیده است (جدول ۴)؛ ذرات سیلت در مقایسه با زبری ذرات شن و چسپندگی ذرات رس از خاصیت نرمی برخوردار هستند (Campbell, 1985) بنابراین انتظار می‌رود با افزایش درصد سیلت در خاک به علت ایجاد حالت سست و نرم، مقاومت فروروی خاک کاهش یابد. در مدل MLR شماره ۳، علاوه بر FWC و سیلت، جرم مخصوص ظاهری نسبی (BDrel) نیز با علامت منفی وارد مدل گردید. به‌نظر می‌رسد با افزایش سهم رس در توزیع اندازه ذرات خاک، جرم مخصوص ظاهری مرجع در مخرج کسر محاسبه BDrel کاهش و در نتیجه BDrel افزایش و منجر به کاهش PR می‌گردد. کلیشادی و همکاران (Kelishadi *et al.*, 2014) توضیح نسبتاً کاملی درخصوص کاستی‌های جرم مخصوص ظاهری و مزیت های استفاده از BDrel به‌جای آن ارائه کرده‌اند. همان‌طوری‌که از جداول ۳ و ۴ نمایان است برخلاف انتظار، در این پژوهش بین کربن آلی و PR همبستگی معنی‌دار یافت نشد بنابراین کربن آلی نتوانست به‌عنوان یک متغیر ورودی مدل‌های رگرسیونی تخمین‌زننده PR ظاهر گردد که دلیل این مسأله را شاید بتوان به ضریب تغییرات بالای آن (CV= 60%) در سه کاربری مجاور هم جنگلی، مرتعی و زارعی منطقه فندقلوی اردبیل نسبت داد. این در حالی است که در تحقیق شیخزاده و همکاران (Sheykhzadeh et al., 2016) در خاک‌های دشت اردبیل، بین کربن آلی و PR همبستگی منفی و معنی‌دار (r= −۰/۴۷**) پیدا شد و کربن آلی در برخی مدل‌های MLR برای برآورد PR، به‌عنوان ورودی مدل انتخاب گردید.

جدول ۳- ضریب همبستگی پیرسون (r) بین متغیرهای مورد مطالعه (n=80)

Table 3- Pearson correlation coefficient (r) between studied variables (n=80)

Variable	OC	Sand	Silt	Clay	dg	σ_g	FWC	BDrel	PD	MWDd	GMD	$CaCO_3$
PR	-0.06	0.21	-0.36**	0.14	0.10	0.36**	-0.45**	0.17	0.19	0.08	-0.01	-0.03
OC	1	0.59**	-0.24*	-0.59**	0.55**	0.03	0.78**	-0.50**	-0.57**	0.23*	0.14	0.40**
Sand		1	-0.76**	-0.56**	0.88**	0.42**	0.44**	-0.19	-0.32*	0.18	0.11	0.42**
Silt			1	-0.11	-0.50**	-0.84**	-0.07	-0.17	0.07	-0.24*	-0.27*	-0.28*
Clay				1	-0.71**	0.43**	-0.59**	0.50**	0.15	0.04	0.18	-0.29*
dg					1	0.01	0.47**	-0.34**	-0.36**	0.02	-0.03	0.45**
σ_g						1	-0.12	0.33**	0.08	0.34**	0.36**	0.05
FWC							1	-0.57**	-0.63**	0.25*	0.17	0.37**
BDrel								1	0.52**	-0.03	-0.01	-0.19
PD									1	-0.14	-0.08	-0.28*
MWDd										1	0.87**	-0.07
GMD											1	0.02
$CaCO_3$												1

*و **: به‌ترتیب معنی دار در سطح احتمال ۱ و ۵ درصد. توصیف علائم متغیرها در زیر جدول ۲ آورده شده است.

Variable symbols description is available under Table 2. *, **: Significant at $P < 0.05$ and $P < 0.01$, respectively.

جدول ٤- مدل‌های رگرسیون خطی چندگانه (MLR) مقاومت فروروی (PR) خاک

Table 4- MLR models of soil PR

شماره Number	مدل Model	حداکثر VIF Max VIF
1	PR= 3.007 - 0.035 FWC	1.00
2	PR= 4.22 - 0.041 FWC - 0.024 silt	1.03
3	0.059 FWC - 0.030 silt – 5.098 BDrel PR= 8.161-	1.40

توصیف متغیرها در زیر جدول ۲ آورده شده است.

Variables description is available under Table 2.

جدول ۵- معیارهای ارزیابی مدل‌های رگرسیونی خطی چندگانه (MLR) مقاومت فروروی (PR) خاک
Table 5- Evaluation criteria for MLR models of PR

شماره Number	متغیرهای ورودی Input variables	آزمون Testing set				آموزش Training			
		R^2	RMSE (MPa)	ME (MPa)	NS	R^2	RMSE (MPa)	ME (MPa)	NS
1	FWC	0.27	1.32	0.12	0.20	0.18	0.77	-0.001	0.18
2	FWC, silt	0.44	1.19	0.19	0.36	0.31	0.70	0.003	0.31
3	FWC, silt, BDrel	0.34	1.25	0.19	0.29	0.45	0.63	0	0.45

NS: ضریب نش-ساتکلیف، RMSE: مجذور میانگین مربعات خطا، ME: میانگین خطا، R^2: ضریب تبیین. توصیف متغیرها در زیر جدول ۲ آورده شده است.
R^2: coefficient of determination; RMSE: root mean square error; ME: mean error; NS: Nash-Sutcliff coefficient. Variables description is available under Table 2.

در جدول ۵ مقادیر معیارهای ضریب تبیین (R^2)، مجذور میانگین مربعات خطا (RMSE)، میانگین خطا (ME) و ضریب نش- ساتکلیف (NS) برای ارزیابی مدل‌های MLR از نظر دقت برآورد PR آورده شده است. استفاده از R^2 به تنهایی برای سنجش اعتبار مدل‌ها ممکن است فریبنده باشد و الزاماً R^2 نزدیک به ۱ دلیل بر تطابق کامل مقادیر اندازه‌گیری با مقادیر برآوردی PR نیست یعنی بالا بودن R^2 دلیلی بر کم بودن خطا نیست (Kozak *et al.*, 1996) بنابراین در پژوهش حاضر رتبه‌بندی مدل‌های MLR، ANN و GEP در برآورد PR، بر مبنای پایین بودن RMSE و بالا بودن NS انجام شد. براساس پایین بودن میزان RMSE و بالا بودن میزان NS در داده‌های آزمونی، مشاهده می‌شود مدل MLR شماره ۲ (جدول ۵) با متغیرهای ورودی رطوبت خاک مزرعه و سیلت از نظر برآورد PR دارای دقت بالاتری در مقایسه با سایر مدل‌های MLR بود. مقادیر R^2، RMSE، ME و NS این مدل برای داده‌های آموزشی به‌ترتیب برابر ۰/۳۱، ۰/۷۰ MPa، ۰/۰۰۳ MPa و ۰/۳۱ و برای داده‌های آزمونی برابر ۰/۴۴، ۱/۱۹ MPa، ۰/۱۹ MPa و ۰/۳۶ تعیین گردید (جدول ۵). مقادیر مثبت ME برای داده‌های آموزشی و آزمونی بیان‌گر برآورد بیشتر از مقدار واقعی PR خاک توسط مدل شماره ۲ می‌باشد. همچنین معادله رگرسیونی این مدل (جدول ۴) نشان می‌دهد که تأثیر منفی متغیر رطوبت خاک مزرعه به علت داشتن ضریب رگرسیونی بالاتر در برآورد PR به مراتب بیشتر از متغیر سیلت می‌باشد. در رتبه‌های بعدی، مدل شماره ۳ با متغیرهای ورودی رطوبت خاک مزرعه، سیلت و جرم مخصوص ظاهری نسبی و مدل ۱ با متغیر ورودی رطوبت خاک مزرعه قرار دارد (جدول ۵)؛ پس در مجموع براساس آماره‌های ارزیابی سه مدل رگرسیونی فوق‌الذکر، مدل شماره ۲ برای برآورد PR ارجح‌تر است. در پژوهش شیخزاده و همکاران (Sheykhzadeh *et al.*, 2016) نیز بهترین مدل MLR با متغیرهای ورودی رطوبت خاک مزرعه و تخلخل کل با مقادیر R^2، RMSE و ME به‌ترتیب برابر با ۰/۴۱، ۱/۱۶ MPa، ۰/۱۷ MPa (داده‌های آزمونی) برای برآورد PR در خاک‌های زراعی دشت اردبیل تعیین گردید.

نتایج مدل‌های شبکه عصبی مصنوعی (ANN) برای برآورد مقاومت (PR) خاک

برای ایجاد مدل‌های شبکه عصبی مصنوعی (ANN) برآورد کننده PR، از ۳ مدل با همان متغیرهای ورودی مدل‌های MLR (جدول ۵) استفاده شد؛ جدول ۶ معیارهای ارزیابی ۳ مدل ANN ارائه شده برای برآورد PR را نشان می‌دهد. رتبه‌بندی مدل‌های ANN نیز همانند مدل‌های MLR (جدول ۵)، بر مبنای پایین بودن RMSE و بالا بودن NS انجام شد. براساس معیارهای فوق در داده‌های آزمونی (جدول ۶)، بهترین مدل ANN برای برآورد PR، مدل ANN شماره ۳ با متغیرهای ورودی رطوبت خاک مزرعه، سیلت و جرم مخصوص ظاهری نسبی در مقایسه با سایر مدل‌های ANN بود. مقادیر R^2، RMSE، ME و NS این مدل برای داده‌های آموزشی به‌ترتیب برابر ۰/۷۹، ۰/۳۹ MPa، ۰/۰۵ MPa و ۰/۷۸ و برای داده‌های آزمونی برابر ۰/۹۲، ۰/۴۱ MPa، ۰/۰۵ MPa – و ۰/۹۲ تعیین گردید (جدول ۶). مقادیر مثبت و منفی ME برای داده‌های آموزشی و آزمونی به‌ترتیب بیان‌گر برآورد جزئی بیشتر و کمتر از مقدار واقعی PR خاک توسط مدل شماره ۳ می‌باشد. در رتبه‌های بعدی، به‌ترتیب مدل ANN شماره ۲ و شماره ۱ دارای دقت بیشتری در برآورد PR می‌باشد؛ مدل ANN شماره ۲ نشان می دهد اضافه شدن متغیر سیلت به مدل ANN شماره ۱ موجب کاهش خطای مدل و افزایش دقت آن گردید به‌طوری‌که براساس داده‌های آزمونی، RMSE مدل ANN شماره ۲ نسبت به مدل ANN شماره ۱ به میزان ۴۵ درصد کاهش و ضریب NS مدل ANN شماره ۲ نسبت به مدل ANN شماره ۱ به میزان ۹۳ درصد افزایش یافت. بیات و همکاران (Bayat *et al.*, 2008) گزارش کردند افزودن متغیر درجه اشباع نسبی به مدل ANN برای برآورد PR، موجب بهبود تخمین مقاومت فروروی خاک از طریق افزایش R^2 از ۰/۷۵ به ۰/۸۰ و کاهش RMSE به میزان ۸/۶۵ درصد گردید. سانتوس و همکاران (Santos *et al.*, 2011) بیان کردند که مقاومت فروروی خاک با دقت بالایی (RMSE= ۰/۰۸۴ و R^2= ۰/۹۸) توسط شبکه عصبی مصنوعی از روی متغیرهای ورودی جرم مخصوص ظاهری و رطوبت خاک مزرعه مورد

تخمین قرار گرفت. شیخزاده و همکاران (Sheykhzadeh et al., 2016) نیز در خاک‌های زراعی دشت اردبیل، مدل با متغیرهای ورودی رطوبت خاک مزرعه، سیلت، جرم مخصوص ظاهری و انحراف استاندارد هندسی (σ_g) ذرات خاک را به‌عنوان بهترین مدل ANN (R^2، RMSE و ME به‌ترتیب برابر با ۰/۴۱، ۱/۴ MPa، ۰/۲۰۶ MPa – در داده‌های آزمونی) برای برآورد PR خاک معرفی کردند.

نتایج مدل‌های برنامه‌ریزی بیان ژن (ANN) برای برآورد مقاومت فروروی (PR) خاک

در برآورد PR با استفاده برنامه‌ریزی بیان ژن (GEP) نیز ۳ نوع مدل با همان متغیرهای ورودی مدل‌های MLR و ANN (جداول ۵ و ۶) ایجاد گردید؛ جدول ۷ معیارهای ارزیابی ۳ مدل GEP ارائه شده برای برآورد PR را نشان می‌دهد. رتبه‌بندی مدل‌های GEP نیز همانند مدل‌های MLR (جدول ۵) و ANN (جدول ۶)، براساس پایین بودن RMSE و بالا بودن NS انجام گردید. بر مبنای آماره‌های مذکور در داده‌های آزمونی (جدول ۷)، دقیق‌ترین مدل GEP برای برآورد PR، مدل شماره ۳ با متغیرهای ورودی رطوبت خاک مزرعه، سیلت و جرم مخصوص ظاهری نسبی در مقایسه با سایر مدل‌های GEP می‌باشد. مقادیر R^2، RMSE، ME و NS برای این مدل براساس داده‌های آموزشی به‌ترتیب برابر ۰/۷۵، ۰/۴۳ MPa، ۰/۰۴ MPa و ۰/۷۴ و داده های آزمونی برابر ۰/۷۹، ۰/۹۱ MPa، ۰/۱۳ MPa و ۰/۶۳ به‌دست آمد (جدول ۷). مقادیر مثبت ME در داده‌های آموزشی و آزمونی نشان‌دهنده اندکی بیش‌برآوردی در تخمین PR خاک توسط مدل GEP شماره ۳ می‌باشد. در رتبه‌های بعدی، به‌ترتیب مدل GEP شماره ۲ و شماره ۱ دارای بیشترین دقت در برآورد PR می‌باشد؛ با توجه به جدول ۷ مشاهده می‌شود افزودن متغیر سیلت به مدل GEP شماره ۱ باعث کاهش RMSE از ۱/۲۶ به ۰/۹۳ مگاپاسکال و افزایش ضریب NS از ۰/۲۸ به ۰/۶۱ در داده‌های آزمونی گردید. دیناروند و همکاران (Dinarvand et al., 2018) نیز بیان کردند مدل GEP با متغیرهای ورودی رطوبت خاک مزرعه، جرم مخصوص ظاهری، رس و میانگین هندسی قطر (d_g) ذرات خاک توانست PR خاک در اراضی زراعی دشت اردبیل با دقت بالایی تخمین بزند به‌طوری‌که مقادیر R^2، RMSE و ME مدل مذکور به‌ترتیب برابر با ۰/۵۱، ۱ MPa و ۰/۴۳- MPa در داده‌های آزمونی به‌دست آمد.

جدول ٦– معیارهای ارزیابی مدل‌های شبکه عصبی مصنوعی مقاومت فروروی (PR) خاک

Table 6 - Evaluation criteria for ANN models of PR

		آزمون Testing set				آموزش Training			
شماره Number	متغیرهای ورودی Input variables	R^2	RMSE (MPa)	ME (MPa)	NS	R^2	RMSE (MPa)	ME (MPa)	NS
1	FWC	0.56	1.12	0.17	0.43	0.14	0.79	0.07	0.13
2	FWC, silt	0.89	0.61	-0.29	0.83	0.70	0.46	-0.03	0.70
3	FWC, silt, BDrel	0.92	0.41	-0.05	0.92	0.79	0.39	0.05	0.78

NS: ضریب نش–ساتکلیف، RMSE: مجذور میانگین مربعات خطا، ME: میانگین خطا، R^2: ضریب تبیین. توصیف متغیرها در زیر جدول ۲ آورده شده است.

R^2: coefficient of determination; RMSE: root mean square error; ME: mean error; NS: Nash-Sutcliff coefficient. Variables description is available under Table 2.

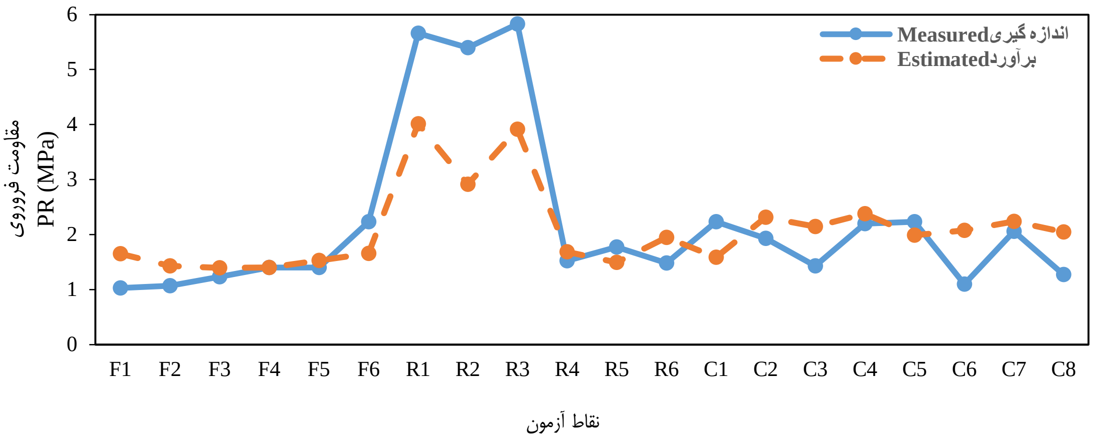
جدول ۷– معیارهای ارزیابی مدل‌های برنامه‌ریزی بیان ژن مقاومت فروروی (PR) خاک

Table 7- Evaluation criteria for GEP models of PR

		آزمون Testing set				آموزش Training			
شماره Number	متغیرهای ورودی Input variables	R^2	RMSE (MPa)	ME (MPa)	NS	R^2	RMSE (MPa)	ME (MPa)	NS
1	FWC	0.30	1.26	0.08	0.28	0.55	0.59	-0.02	0.50
2	FWC, silt	0.74	0.93	0.13	0.61	0.61	0.53	-0.03	0.61
3	FWC, silt, BDrel	0.79	0.91	0.13	0.63	0.75	0.43	0.04	0.74

NS: ضریب نش–ساتکلیف، RMSE: مجذور میانگین مربعات خطا، ME: میانگین خطا، R^2: ضریب تبیین. توصیف متغیرها در زیر جدول ۲ آورده شده است.

R^2: coefficient of determination; RMSE: root mean square error; ME: mean error; NS: Nash-Sutcliff coefficient. Variables description is available under Table 2.



شکل ۳- بررسی عملکرد بهترین مدل برنامه‌ریزی بیان ژن (GEP) در برآورد مقاومت فروروی (PR) خاک در نقاط آزمونی (n=20)؛ F: کاربری جنگلی، R: کاربری مرتعی و C: کاربری زراعی

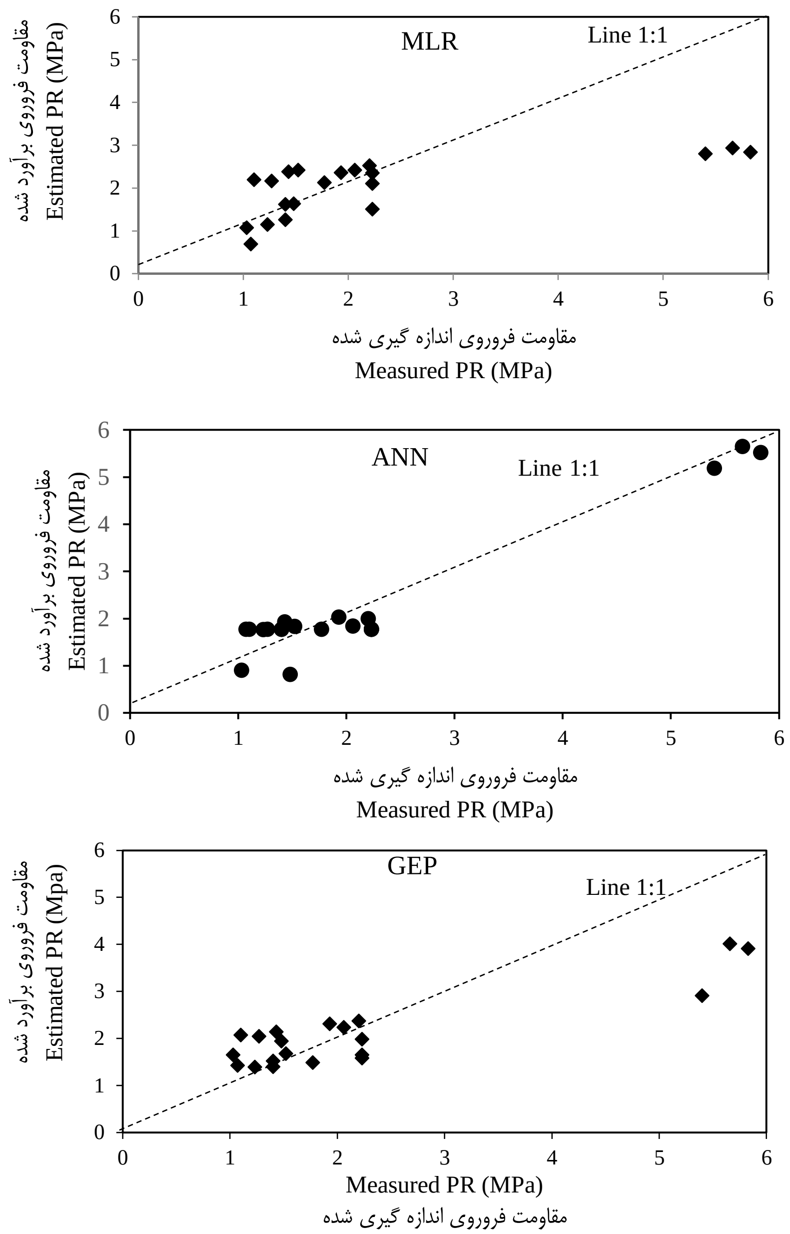
Figure 3- Investigation the performance of GEP model in estimating soil penetration resistance (PR) at the testing points (n=20); F: forest land use, R: range land use and C: cultivated land use

بررسی عملکرد بهترین مدل در برآورد مقاومت فروروی (PR) خاک در نقاط آزمون (۲۰ داده تصادفی) به‌عنوان نمونه در شکل ۳ برای مدل برنامه‌ریزی بیان ژن (GEP) نشان داده شده است؛ با توجه به این شکل مشاهده می‌شود خطای مدل (اختلاف بین مقادیر اندازه‌گیری شده با مقادیر برآورد شده) در پیش‌بینی متغیر وابسته (PR) از روی متغیرهای مستقل (رطوبت خاک مزرعه، سیلت و جرم مخصوص ظاهری نسبی) در نمونه‌های خاک برداشته شده از کاربری مرتعی (R) به مراتب بیشتر از کاربری‌های جنگلی (F) و زراعی (C) می‌باشد؛ علت این مسأله را شاید بتوان به ضریب تغییرات (CV) بالای مقاومت فروروی خاک در کاربری مرتعی (۵۸ درصد) در مقایسه با کاربری‌های جنگلی (۲۲ درصد) و زراعی (۲۳ درصد) نسبت داد همچنین از بین متغیرهای مستقل نیز رطوبت خاک مزرعه دارای بیشترین ضریب تغییرات در کاربری مرتعی (۳۷ درصد) در مقایسه با کاربری‌های جنگلی (۱۶ درصد) و زراعی (۱۶ درصد) بود. نتایج تحقیقات قبلی نیز بیانگر تأثیرپذیری بیشتر مقاومت فروروی از رطوبت خاک مزرعه در مقایسه با سایر ویژگی‌های زودیافت خاک می‌باشد (Bayat *et al.*, 2008; Vaz *et al.*, 2011; Santos *et al.*, 2011; Sheykhzadeh *et al.*, 2016; Asghari & Shahabi, 2019).

مقایسه دقت مدل‌های MLR، ANN و GEP در برآورد PR

مقادیر R^2، RMSE و NS برای مدل‌های MLR (جدول ۵)، ANN (جدول ۶) و GEP (جدول ۷) ارائه شده برای برآورد PR در داده‌های آزمونی بیانگر آن است که مدل‌های شبکه‌های عصبی مصنوعی (ANN) به علت داشتن R^2 بالا (از ۰/۵۶ تا ۰/۹۲) و NS بیشتر (از ۰/۴۳ تا ۰/۹۳) و RMSE پایین (از ۰/۴۱ تا ۱/۱۲ مگاپاسکال) در مقایسه با مدل‌های رگرسیون خطی چندگانه (R^2 از ۰/۲۷ تا ۰/۳۴، NS از ۰/۲۰ تا ۰/۳۶و RMSE از ۱/۲۵ تا ۱/۳۲ مگاپاسکال) و مدل های برنامه‌ریزی بیان ژن (R^2 از ۰/۳۰ تا ۰/۷۹، NS از ۰/۲۸ تا ۰/۶۳و RMSEاز ۰/۹۱ تا ۱/۲۶ مگاپاسکال) دارای دقت بالاتری و خطای کمتری در برآورد PR بودند. شکل ۴ نیز نمودار ۱:۱ مقادیر اندازه‌گیری و برآورد شده PR خاک را با استفاده از بهترین مدل MLR (مدل شماره ۲ در جدول ۵)، بهترین مدل ANN (مدل شماره ۳ در جدول ۶) و بهترین مدل GEP (مدل شماره ۳ در جدول ۷) برای داده‌های آزمونی نشان می‌دهد. با توجه به این شکل مشاهده می‌شود که پراکندگی داده ها حول محور ۱:۱ در بهترین مدل ANN کمتر از مدل‌های MLR و GEP برای برآورد PR در خاک‌های منطقه مورد مطالعه می‌باشد همچنین می‌توان گفت مدل‌های شبکه عصبی مصنوعی و برنامه‌ریزی بیان ژن در تخمین PR بەدلیل استفاده از روابط غیر خطی به مراتب دقیق‌تر از مدل‌های رگرسیونی خطی چندگانه عمل نموده‌اند. علت برآورد دقیق‌تر شبکه‌های عصبی مصنوعی نسبت به مدل‌های رگرسیونی را می‌توان در نحوه عمل شبکه پرسپترون سه لایه دانست. عملکرد این شبکه به این صورت است که ابتدا الگویی به شبکه عصبی عرضه می‌شود و خروجی آن محاسبه می‌شود سپس شبکه این خروجی را با مقادیر اندازه‌گیری شده متغیر یا متغیرهای خروجی مقایسه می‌کند و ضرایب وزنی شبکه را به نوعی تغییر می‌دهد که کمترین اختلاف بین مقادیر پیش‌بینی شده و اندازه‌گیری شده باشد. بیات و همکاران (Bayat *et al.*, 2008) و شیخزاده و همکاران (Sheykhzadeh *et al.*, 2016) نیز گزارش نمودند که شبکه‌های عصبی مصنوعی نسبت به مدل‌های رگرسیون خطی چندگانه در برآورد PR دقیق‌تر عمل نمودند. سانتوس و همکاران (Santos *et al.*, 2011) نشان دادند تابع شبکه عصبی

مصنوعی (R^2= ۰/۹۸ و RMSE= ۰/۰۸۴) نسبت به تابع رگرسیونی (R^2= ۰/۹۲ و RMSE= ۰/۹۵۱) دقت بیش‌تری در برآورد مقاومت فروروی از روی جرم مخصوص ظاهری و رطوبت خاک مزرعه داشت. نتایج مشابهی در خصوص دقت بالای مدل GEP در مقایسه با مدل‌های رگرسیون خطی و غیرخطی در برآورد ویژگی‌های دیریافت خاک از روی ویژگی‌های زودیافت خاک توسط احمدزاده کلیبر و فولاپناه (Ahmadzadeh Kaleibar & Fuladipanah, 2023) و زانگ و زانگ (Zhang & Zhangh, 2024) گزارش گردیده است.



شکل ٤- مقایسه مقادیر اندازه‌گیری و برآورد شده مقاومت فروروی (PR) خاک با استفاده از بهترین مدل رگرسیون خطی چندگانه (MLR)، شبکه عصبی مصنوعی (ANN) و برنامه‌ریزی بیان ژن (GEP) بر اساس داده‌های آزمونی (n=20). اطلاعات مربوط به مدل‌ها و متغیرهای ورودی در جداول ۵، ٦ و ۷ آورده شده است. خطچین بیانگر نمودار ۱:۱ است.

Figure 4- Comparisons of estimated and measured values of penetration resistance (PR) for the testing data set based on the best MLR, ANN and GEP models. Tables 4, 5 and 6 provide detailed information about each model and input variables. The dashed lines indicate the 1:1 diagonals.

نتیجه‌گیری

در این تحقیق از مدل‌های رگرسیونی خطی چندگانه (MLR)، شبکه عصبی مصنوعی (ANN) و برنامه‌ریزی بیان ژن (GEP) برای برآورد مقاومت فروروی (PR) خاک از روی ویژگی‌های زودیافت خاک در اراضی جنگلی، مرتعی و زراعی فندقلوی اردبیل استفاده گردید. نتایج نشان داد در منطقه مورد مطالعه، متغیرهای رطوبت خاک مزرعه، سیلت و جرم مخصوص ظاهری نسبی در برآورد PR در مقایسه با سایر ویژگی‌های زودیافت خاک از اهمیت بیشتری برخوردار بوده و در ورودی مدل‌ها به‌کار گرفته شدند. مقایسه نتایج مدل‌ها نشان داد ANN به علت داشتن ضرایب تبیین (R^2) و نش-ساتکلیف (NS) بالاتر و مجذور میانگین مربعات خطای (RMSE) پایین‌تر، دارای دقت بالاتر و خطای کمتری در برآورد PR خاک نسبت به مدل‌های MLR بودند. پیشنهاد می‌شود در تحقیقات بعدی از روش رگرسیون غیر خطی و نیز سایر مدل‌های هوشمند مانند نروفازی، ماشین بردار پشتیبان و جنگل تصادفی برای برآورد مقاومت فروروی خاک استفاده گردیده و نتایج آن با مدل‌های رگرسیونی مقایسه گردد.

سپاسگزاری

این مقاله مستخرج از طرح پژوهشی مصوب دانشگاه محقق اردبیلی می‌باشد به این وسیله از حمایت‌های معاونت محترم پژوهشی و فن آوری دانشگاه تشکر و قدردانی می‌شود.

Referensec

1- Asghari, Sh., & Shahabi M. (2019). Spatial variability of soil saturated hydraulic conductivity and penetration resistance in salt-affected lands around Lake Urmia. *Water and Soil*, *33*(1), 103-116. (In Persian with English abstract). https://doi.org/10.22067/JSW.V33I1.74411

2- Asghari, Sh., Sheykhzadeh, G.R., & Shahabi, M. (2017). Geostatistical analysis of soil mechanical properties in Ardabil plain of Iran. *Archives of Agronomy and Soil Science*, *63*(12), 1631-1643. https://doi.org/10.1080/03650340.2017.1296136

3- Azadmard, B., Mosaddeghi, M.R., Ayoubi, S., Chavoshi, E., & Raoof, M. (2019). Estimation of near-saturated soil hydraulic properties using hybrid genetic algorithm-artificial neural network. *Ecohydrology & Hydrobiology*, *20*(3), 437–449. https://doi.org/10.1016/j.ecohyd.2019.09.001

4- Ahmadzadeh Kaleibar, F., & Fuladipanah, M. (2023). Assessment of regression, support vector machine, and gene expression programming transfer functions to predict soil humidity parameters in Arasbaran plain. *Journal of Water and Soil Science*, *27*(2), 135-149. https://doi.org/10.47176/jwss.27.2.42532

5- Ahmadi, A., Palizvan zand, P., & Palivan zand, H. (2018). Estimation of saturated hydraulic conductivity by using gene expression programming and ridge regression (A case study in East Azerbaijan province). *Iranian Journal of Soil and Water Research*, *48*(5), 1087-1095. (In Persian with English abstract). https://doi.org/10.22059/ijswr.2018.218413.667555

6- Amirabedi, H., Asghari, Sh., Mesri, T., & Balandeh, N. (2016). Prediction of mean weight diameter of aggregates using artificial neural network and regression models. *Applied Soil Research*, *4*(1), 39-53. (In Persian with English abstract)

7- Bayat, H., Neyshabouri, M.R., & Hajabbasi, M. (2008). Comparing neural networks, linear and nonlinear regression techniques to model penetration resistance. *Turkish Journal of Agriculture and Forestry*, *32*, 425-433.

8- Blake, G.R., & Hartge, K.H. (1986a). *Bulk density*. p. 363-375. In: Klute A. (ed). Methods of Soil Analysis Part 1, Physical and Mineralogical Methods. 2nd ed. American Society of Agronomy, Madison, WI.

9- Blake, G.R., & Hartge, K.H. (1986b). *Particle density*. p. 377-381. In: Klute A. (ed). Methods of Soil Analysis Part 1, Physical and Mineralogical Methods. 2nd ed. ASA and SSSA, Madison, WI.

10- Campbell. G.S. (1985). *Soil Physics with Basic: Transport Models for Soil–Plant System*. Elsevier. New York. 150 p.

11- Dinarvand, H., Asghari, Sh., & Shahabi, M. (2018). Estimating soil penetration resistance using neurofuzzy, support vector machine and gene expression programming methods. Available at https://civilica.com/doc/808626 (In Persian with English abstract)

12- Ferreira, C. (2001). Gene expression programming: a new adaptive algorithm for solving problems. *Complex Syst 13*, 87–129.

13- Ghorbani, M.A., Deo, R.C., Kashani, M.H., Shahabi M., & Ghorbani, S. (2019). Artificial intelligence-based fast and efficient hybrid approach for spatial modeling of soil electrical conductivity. *Soil and Tillage Research*, *186*, 152–164. https://doi.org/10.1016/j.still.2018.09.012

14- Gardner, W.H. (1986). *Water content*. p. 493-544. In: Klute A. (ed). Methods of Soil Analysis. Part 1. 2nd ed. Agronomy. Monograph. 9. ASA, Madison, WI.

15- Gee, G.W., & Or, D. (2002). *Particle-size analysis*. p. 255–293. In: Dane J. H., & Topp G. C. (eds.). Methods of Soil

Analysis. Part 4. SSSA Book Series No. 5. Soil Science Society of America, Madison, WI.

16- Kelishadi, H., Mossaddeghi, M.R., Hajabbasi, M.A., & Ayoubi, S. (2014). Near-saturated soil hydraulic properties as influenced by land use management systems in Koohrang region of central Zagros, Iran. *Geoderma*, *213*, 426-434. https://doi.org/10.1016/j.geoderma.2013.08.008

17- Kozak, E., Pachepsky, Y.A., Sokolowski, S., Sokolowska, Z., & Stepniewski, W. (1996). A modified number-based method for estimating fragmentation fractal dimensions of soils. *Soil Science Society of America Journal*, *60*, 1291-1297.

18- Lowery, B., & Morrison, JE. (2002). *Soil penetrometer and penetrability*. In: Dane J.H., & Topp GC (eds.). Methods of soil analysis, part 4. Physical methods. Madison (WI): Soil Science Society of America; pp. 363–388.

19- Nelson, D.W., & Sommers, L.E. (1982). *Total carbon, organic carbon, and organic matter*. p. 539–579. In A.L. Page et al. (ed.) Methods of Soil Analysis. Part 2. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI.

20- Page, A.L. (ed.).(1985). *Methods of Soil Analysis. Part 2. Chemical and Microbiological Methods*. Agronomy No. 9. American Society of Agronomy, Madison, WI.

21- Santos, F.L., De Jesus, V.A.M., & Valente, D.S.M. (2012). Modeling of soil penetration resistance using statistical analyses and artificial neural networks. *Acta Scientiarum. Agronomy*, *34*, 219-224.

22- Sheykhzadeh, G.R., Asghari, Sh., & Mesri Gundoshmian, T. (2016). Estimating penetration resistance in agricultural soils of Ardabil plain using artificial neural network and regression methods. *Water and Soil*, *30*(3), 941-954. (In Persian with English abstract). https://doi.org/10.22067/JSW.V30I3.42235

23- Vaz, C.M.P., Manieri, J.M., de Maria, I.C., & Tuller, M. (2011). Modeling and correction of soil penetration resistance for varying soil water content. *Geoderma*, *166*, 92-101. https://doi.org/10.1016/j.geoderma.2011.07.016

24- Whalley, W., To, J., Kay, B.D., & Whitmore, A.P. (2007). Prediction of the penetrometer resistance of soils with models with few parameters. *Geoderma*, *137*, 370–377. https://doi.org/10.1016/j.geoderma.2006.08.029Get rights and content

25- Yazdani, A., Mosaddeghi, M.R., Khademi, H., Ayoubi, S., & Khayamim, F. (2014). Relationship between surface aggregate stability and some soil and climate properties in Isfahan province. *Soil Management*, *3*(2), 23-31. (In Persian with English abstract)

26- Yoder, R.E. (1936). A direct method of aggregate analysis of soils and a study of the physical nature of erosion losses. *Journal of American Society Agronomy*, *28*, 337-35.

27- Zhang, R., & Zhang, S. (2024). Coefficient of permeability prediction of soils using gene expression programming. *Engineering Applications of Artificial Intelligence*, *128*(107504). https://doi.org/10.1016/j.engappai.2023.107504